DOCUMENT RESUME

ED 151 382 TM 006 800

AUTHOR Cahen, Leonard S.
TITLE Report of "B" Testing. Beginning Teacher Evaluation Study. Technical Note Series. Technical Note III-3.
INSTITUTION Far West Lab. for Educational Research and Development, San Francisco, Calif.
PUB DATE 25 Feb 77
NOTE 11p.; For related documents, see TM 006 799 and 801

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Tests; Affective Tests; Complexity Level; Elementary Education; *Elementary School Mathematics; Grade 2; Grade 5; Item Analysis; *Reading Tests; *Statistical Data; Student Testing; Teacher Evaluation; *Testing; *Testing Programs; Test Interpretation

ABSTRACT

The "B" testing program of the Beginning Teacher Evaluation Study, Phase III-B, is summarized. A sample of 146 second grade and 128 fifth grade students completed the reading and mathematics achievement battery, as well as affective scales pertaining to reading, mathematics, and school. The administration procedures, test results, and psychometric properties of the tests are described. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151382

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BTES

BEGINNING TEACHER EVALUATION STUDY

Technical Note Series

Technical Note III-3

REPORT OF "B" TESTING

Leonard S. Cahen

February 25, 1977

FAR WEST LABORATORY
FOR EDUCATIONAL RESEARCH AND DEVELOPMENT
1855 FOLSOM STREET • SAN FRANCISCO CALIFORNIA 94103 • (415) 565-3000

TM 006800

TECHNICAL NOTE REPORT OF "B" TESTING

Overview

This report describes the "B" testing. The second grade "target students" were tested between December 6, 1976 and January 7, 1977. The fifth grade "target students" were tested between December 7, 1976 and January 14, 1977. These periods include the regularly scheduled testing dates plus make-up testing for the "target students" who were absent for the regularly scheduled test administrations.

One hundred forty-six second grade target students were tested. The students came from 25 classrooms. Twenty-seven teachers were involved with these students. The number of teachers is greater than the number of classes tested because some classes are team taught. Four second grade target students were lost from the time of the administration of the "A" test battery. At present we have 22 classes with six target students, two classes with five target students, and one class with four target students. The four missing students moved out of the participating schools.

At fifth grade we tested 128 target students which was a decrease of six from the time of the administration of the "A" test battery. One class with six students was dropped from the study. The 128 target students come from 21 classes, taught by a total of 29 teachers. At present we have 19 classes with six target students, one team taught class with nine target students, and one class with five target students.

Test Administration

The test administrators were again given one-half day of training to familiarize themselves with the test booklets, directions, time limits, etc. The training was designed to help them prepare for administering

tests to a small number of pupils. Again, special rooms were set up in each school for the administration of the tests. Target students were removed from their regular classrooms and tested as a group. After testing, the booklets were returned to the Far West Laboratory for processing. Absent students were rescheduled for testing. The test administrators reported that, in general, the testing situations went well and were similar to procedures and practices that occurred in conjunction with the "A" testing (see Technical Note III-2, Cahen, 1976)

Data Processing and the Return of Information to Teachers

All the collected test data were run through an item analysis program which provided a check on the scoring procedures, keys, etc. Psychometric properties, including the total score distributions, were examined for each of the subscales. Using the format that was developed for the reporting of "A" scores to teachers, a computerized printout for each teacher, listing information about "B" testing for each of their target students, was prepared. Each teacher was mailed a copy of the scores for his/her students plus a Manual of Interpretation. These materials were sent to the teachers by Certified Mail (return receipt requested). This procedure was again used to maintain our policy of confidentiality of data by having the test scores returned directly to the appropriate teacher. We have continued to meet our obligation to teachers after having promised them that we would share our data with them.

Psychometric Properties of the Tests

Tables 1 and 2 display a summary of psychometric data for the Grade 2 Reading and Mathematics subtests. Tables 3 and 4 display a summary psychometric data for the Grade 5 Reading and Mathematics Subtests. All scores are corrected for chance guessing.

Insert Tables 1 - 4 Here

In Tables 1 through 4 the symbol $\overline{P}$ serves as an index of the overall difficulty of the tests. This statistic is defined as the subtest mean (corrected for chance) divided by the number of items. The reader will observe a wide range of alpha (reliability) coefficients in Tables 1 - 4.

The last column in Tables 1 - 4 shows the percent of students passing 80% or more of the items on the subtest. We are again concerned about the possibility of "ceiling effects" which will limit the amount of positive growth or change we can observe for the target students. We are presently in the process of reviewing our testing plans for the Spring "C" test administration and the data from the last column in Tables 1 - 4 will be carefully attended to by BTES staff.

Table 5 shows the summary psychometrics of the three Affective Scales, Grades 2 and 5. The reader is reminded that, for Grade 2, each item in the scale can earn a score of 1 - 3 points. For Grade 5, the "Likert-type" items have five response categories. The possible range of scores on the reading affect scale for Grade 2 would be from 7 to 21, and from 7 to 35 at Grade 5. The Affect Scales, at both grade levels, appear to be doing well from a psychometric point of view.

Insert Table 5 Here

In general, the staff feels that the "B" testing processs proceeded smoothly.

Table 1

Psychometric Data
Grade 2 Reading Subtests
"B" Administration
(N = 146)

Subtests	# of Items	$\overline{X}$	σ	$\overline{P}$	Alpha	Standard Error of Measurement	Percent passing 80% or more of Items in Subtest
Decoding Consonant Blends	12	8.26	3.17	.69	.64	1.78	40
Decoding Variant Consonants	14	5.47	5.42	.39	.73	2.73	18
Decoding Long Vowels	20	10.60	6.26	.53	.82	2.60	21
Decoding Complex Patterns	16	6.20	4.62	.39	.75	2.26	12
Compound Words	8	5.51	2.09	.69	.76	0.96	40
Root Words and Affixes - D	9	4.74	1.84	.53	.52	1.20	9
Root Words and Affixes - C	8	2.91	2.70	.36	.55	1.71	16
Syllables	10	2.93	2.69	.29	.40	1.97	4
Synonyms	16	5.13	4.96	.32	.81	2.07	6
Pronoun Reference	10	3.26	3.36	.33	.63	1.95	12
Events	11	3.70	3.48	.34	.65	1.95	10
Relationships	6	1.36	2.20	.23	.48	1.45	7
Main Idea	6	1.67	1.89	.28	.25	1.49	5

Table 2

Psychometric Data
Grade 2 Mathematics Subtests
"B" Administration
(N = 146)

Subtests	# of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent passing 80% or more of Items in Subtest
Money	10	5.23	2.99	.52	.68	1.59	23
Linear Measurement	11	7.29	1.86	.66	.29	1.50	21
Addition No Regrouping	16	10.21	4.49	.64	.83	1.82	36
Addition with Regrouping	10	-0.26	3.04	.00	.78	1.29	2
Speed Test - Addition	14	6.67	3.72	.47	.78	1.70	11
Place Value/ Numerals	25	7.88	5.53	.32	.75	2.71	3
Subtraction no Regrouping	16	7.95	5.09	.50	.82	2.08	25
Word Problems	15	4.30	3.14	.29	.48	2.19	4
Computational Transfer	22	6.25	4.94	.28	.68	2.73	1
Speed Test - Subtraction	14	4.76	3.53	.34	.72	1.80	3
Fractions	4	1.04	1.20	.26	.00	1.05	3

Table 3

Psychometric Data
Grade 5 Reading Subtests
"B" Administration
(N = 127)

Subtests	# of Items	$\bar{X}$	σ	$\bar{P}$	Alpha	Standard Error of Measurement	Percent passing 80% or more of Items in Subtest
Syllables	21	9.46	5.95	.45	.77	2.78	7
Root Words and Affixes	15	5.49	3.17	.37	.46	2.25	3
Synonyms	15	6.84	3.75	.46	.66	2.12	11
Pronouns	16	8.55	4.40	.53	.75	2.14	24
Word Meaning in Context	20	7.60	6.95	.38	.87	2.45	16
Comprehension - Literal	14	5.72	4.11	.41	.71	2.12	13
Comprehension - Inferential	9	3.95	2.76	.44	.62	1.61	14
Comprehension - Main Idea - Literal	5	2.18	1.94	.44	.56	1.16	13
Comprehension - Main Idea - Inference	4	1.32	1.43	.33	.28	1.05	5
Topic Sentence	10	3.96	3.32	.40	.70	1.72	11
Fact and Opinion	10	1.75	3.08	.17	.64	1.76	2

Table 4

Psychometric Data
Grade 5 Mathematics Subtests
"B" Administration
(N = 127)

Subtests	# of Items	$\overline{X}$	σ	$\overline{P}$	Alpha	Standard Error of Measurement	Percent passing 80% or more of Items in Subtest
Addition and Subtraction	10	7.28	2.90	.72	.76	1.35	52
Open Multiplication	4	2.50	1.53	.63	.81	.57	38
Properties and Factors	18	12.94	3.00	.72	.78	1.36	32
Number Patterns	8	2.29	2.29	.29	.48	1.54	7
Multiplication	16	9.60	5.68	.60	.89	1.84	41
Speed Test - Multiplication	42	17.73	8.90	.42	.94	2.18	6
Place Value/ Numerals	10	5.13	2.61	.51	.58	1.60	19
Word Problems	12	1.99	3.21	.16	.62	1.88	1
Perimeters	5	1.29	1.72	.26	.50	1.09	4
Area	5	.33	1.29	.07	.26	.99	0
Lines and Figures	6	.96	1.72	.16	.36	1.26	2
Number Pairs	6	.92	1.86	.15	.47	1.23	1
Fractions	15	1.96	3.97	.13	.77	1.85	2
Division	16	4.22	5.07	.26	.82	2.09	9

Table 5

Psychometric Data
Affect Measures
"B" Administration

Grade 2 (N = 146)

Affect Scale	# of Items	$\bar{X}$	σ	Alpha	Standard Error of Measurement
Reading	7	15.87	4.03	.82	1.60
Mathematics	7	16.13	4.41	.88	1.44
School	2	4.76	1.49	.76	.51

Grade 5 (N = 127)

Affect Scale	# of Items	$\bar{X}$	σ	Alpha	Standard Error of Measurement
Reading	11	40.47	8.67	.83	2.80
Mathematics	11	38.68	10.46	.93	2.72
School	2	7.43	2.21	.73	.82

Bibliography

Cahen, Leonard S. Report on "A" Testing, Technical Note III-2. Beginning Teacher Evaluation Study. San Francisco, California: Far West Laboratory for Educational Research and Development, 1976.